涂志芳，徐慧芳.学术图书馆研究数据服务的多元路径[J].图书情报工作,2018,62(22):45－54.

学术图书馆研究数据服务的多元路径*

■ 涂志芳　徐慧芳

中国科学院文献情报中心　北京 100190

摘要：[**目的/意义**]对学术图书馆现有的不同数据服务路径进行研究，为国内其他图书馆探索数据服务提供路径与模式的参考。[**方法/过程**]以前期对14所学术型专业图书馆综合服务、15所高校图书馆数字学术服务、3所加拿大高校图书馆数据服务的调研为基础，总结归纳图书馆数据服务的5种主要路径：数据平台与数据服务、"一站式"资源发现系统与数据发现、整合型数字学术空间与数据服务、开放数据资源导航与数据服务、综合性数据咨询与支持服务，分别围绕5种路径进行典型案例研究并思考需要进一步研究的问题及解决方案。[**结果/结论**]不同服务路径各有优劣及适用性条件，彼此之间既可独立存在、又可协同发挥作用；5种路径都有其存在的价值和必要性，任意一种数据服务路径的存在都有助于构建完整的数据服务体系，帮助图书馆在与出版商、信息技术与服务机构等的竞争与合作中更好地占据数据服务的"领地"，把握数据服务的主动权。

关键词：学术图书馆　研究数据服务　数据管理平台　资源发现系统　数字学术

分类号：G250

DOI:10.13266/j.issn.0252－3116.2018.22.005

1　文献回顾

国内外图书馆研究数据或科学数据服务相关研究渐趋丰富，主要可分为两类：一是聚焦于科学数据管理的研究；二是从更广泛视角进行的数据服务理论与实践研究。整体而言，国外以实践层面的数据服务探索为主，辅之以相关讨论和研究；国内则以理论层面的数据服务探索和对国外数据服务实践研究居多，数据服务实践在少部分图书馆中进行。

聚焦于研究数据管理的研究成果相对丰富，前期阶段主要包括数据管理政策、数据管理工具与系统、数据管理生命周期、科学数据元数据、自然科学与人文社会科学等不同领域研究数据管理等相对宏观的研究内容[1]；研究后期阶段至近年来则逐渐深入到数据出版、数据评审、数据引用等更为具体且微观的方面。从更广泛视角进行的数据服务理论与实践研究较为分散而多元，学者们对数据服务的必要性与重要性、数据服务模型与生态圈、数据服务内容与实施路径、数据服务馆员素养与技能、数据服务实践案例等进行了诸多的分析与思考。

国外学者的数据服务理论研究为实践及实践研究奠定了较好的基础，如M. S. Nelson[2]认为联结数据是迈向图书馆新兴服务领域的第一步，联结数据要求图书馆员更新旧技能并学习新技能，及时捕获用户需求，熟练掌握相关工具来识别研究人员和学生的需求，根据不同需求提供不同服务。更明显的是，国外图书馆界比较注重对图书馆数据服务实践现状的调查反馈，如2012年6月美国大学与研究图书馆协会(Association of College & Research Library, ACRL)对美国、加拿大的学术图书馆进行调查并发布《学术图书馆与研究数据服务：当前实践与未来计划》[3]，2016年12月欧洲研究图书馆协会(LIBER)对馆长进行调查并发布《欧洲学术图书馆研究数据服务》[4]，根据调查可知：①已经提供或计划提供研究数据服务的图书馆越来越多，图书馆高度认同研究数据服务的重要性，也尽可能为数据服务馆员提供专业培训、职业发展的机会；②图书馆提供的研究数据服务以咨询类居多，技术/实践类较少；③图书馆与其他校内单元、研究机构合作提供研究数据服务；④图书馆应积极争取数据服务相关的机会和主动权(如制定研究数据政策)，也需要跨越学科

* 本文系北京市社会科学基金项目"学术图书馆参与数据出版质量控制研究"(项目编号:17XCB007)研究成果之一。

作者简介：涂志芳(ORCID：0000－0002－8323－5205)，助理馆员，硕士；徐慧芳(ORCID：0000－0001－7732－4239)，副研究馆员，硕士，通讯作者，E-mail：xuhf@mail.las.ac.cn。

收稿日期:2018－03－16　修回日期:2018－06－04　本文起止页码:45－54　本文责任编辑:王传清

边界提供更多元的服务并服务更广泛的研究群体。

国内学者对数据服务进行了多维度的研究，其中多为理论性的思考与探索，具有较强的指导意义和启发作用。主要包括：①较早期的研究者把数据服务视为图书馆服务（尤其是参考咨询）的新领域加以讨论[5-6]；②对北美、欧洲等国外学术图书馆数据服务的整体实践进展[7-9]、实践个案[10]或数据服务的某一方面（如元数据[11]）进行研究；③对图书馆数据服务模型、生态圈等的理论性探索[12-13]，理论基础与基础设施、数据资源与服务能力、技术支撑与应用平台等是数据服务体系必不可少的要素；④对图书馆数据服务内容与实施路径的探索，数据服务包括数据开发与推广、检索与发现、管理与监护、存储与关联、咨询与分析、技术支持等内容[14-16]，可采取“集成展现”“嵌入学科”等的服务方式；⑤对数据服务馆员服务模式与技能培训的关注，包括数据馆员岗位设置、数据素养及其培养机制等[17-18]。

综上所述，国内外学者已经对图书馆的研究数据（本研究所指“研究数据”包括人文科学、社会科学与自然科学数据）服务积累了一定的研究成果。立足现有研究基础，本研究欲在更大范围内了解图书馆数据服务及其在图书馆整体服务中的状况，从而更加系统而全面地呈现图书馆数据服务的内容与形式，以期为各类不同图书馆选择各自适合的数据服务路径提供参考和选择。因此，笔者以所调研主题的范围“从大到小”的思路进行3轮调研，总结目前图书馆数据服务常见的5种路径，分别对5种路径进行阐述例证（案例包括但不限于3轮调研中的图书馆）和对比分析，从而梳理出不同路径的内容、特点及其不同的适用性。

2 作为研究基础与线索的调研

2017年7月至12月期间，笔者对14所学术型专业图书馆的综合实践、15所高校图书馆的数字学术服务实践、3所加拿大高校图书馆的数据服务实践进行了3轮调研。其中，专业图书馆调研的主要目的是了解对标图书馆在数字环境下的生存、发展及转型、变革；数字学术服务调研的主要目的是了解图书馆数字学术服务进展及不同的数字学术服务内容；数据服务调研的目的是进一步了解图书馆开展数据服务的路径与内容。

2.1 14所学术型专业图书馆综合实践调研

考虑各国家或地区的科技发展水平、图书馆事业发展水平，本研究初步选择美国、英国、加拿大、澳大利亚、德国、日本、俄罗斯、中国台湾地区等为对象，再综合考虑图书馆地域覆盖范围的平衡性及图书馆的代表性等因素，经过试验性调研之后，最终确定以德国国家科技图书馆[19]、德国国家医学图书馆[20]、美国国立医学图书馆[21]、美国国家农业图书馆[22]、美国国家标准与技术研究院图书馆[23]、加拿大联邦科学图书馆[24]、加拿大国家科学图书馆[25]、澳大利亚联邦科学与工业研究组织图书馆[26]、澳大利亚地球科学局图书馆[27]、日本国立情报学研究所图书馆[28]、日本原子能研究开发机构图书馆[29]、俄罗斯国家公共科技图书馆[30]、英国国家物理实验室图书馆[31]、中国台湾地区“中央研究院”图书馆[32]等14所学术型专业图书馆为调研对象（调研时间为2017年7月至12月），通过访问图书馆网站、图书馆学（协）会网站、查找相关文献与网络资料，了解图书馆所在母体机构和（或）专业图书馆概况、图书馆发展战略与行动计划、图书馆资源建设、图书馆用户服务、研究与发展活动、面临压力与挑战等内容。这14所专业图书馆可分为4类，即国家级科学/专业图书馆、国家级研究机构内部大中型专业图书馆、国家级及以下研究机构内部中小型专业图书馆、专业图书馆虚拟联合体，调研对象选择具有较强的代表性。调研发现，专业图书馆的数据服务主要包括：①数据标识符服务（主要为DOI）：如德国国家科技图书馆、德国国家医学图书馆、加拿大国家科学图书馆等的DOI注册服务；②数据管理与咨询：包括数据管理平台建设与维护、数据咨询与支持、相关培训等；③数据存储：以数字存储库、数据知识库等方式进行研究数据的存储，如德国国家科技图书馆承担数据存储图书馆功能并运营着数字保存系统；④数据出版：包括数据出版平台运营、数据存储与出版、数据出版物索引、数据集下载引用，如德国国家医学图书馆的数字保存系统获得数据认可印章（Data Seal of Approval，DSA）的认证、美国国立医学图书馆数据批量下载、美国国家农业图书馆5个数据集成平台、加拿大联邦科学图书馆数据集检索、澳大利亚联邦科学与工业研究组织图书馆数据门户；⑤数据可视化：包括数据可视化工具、数据可视化培训、数据可视化成果展示等，如美国国家标准与技术研究院图书馆为研究人员提供数据可视化墙的使用、研究成果影响力的可视化分析。

2.2 15所高校图书馆数字学术服务调研

以现有文献[33-35]对图书馆数字学术服务的调查揭示为线索，对若干所学术图书馆进行探索性预调研，最终选取香港中文大学图书馆[36]、华盛顿大学图书

馆[37]、布朗大学图书馆[38]、里士满大学图书馆[39]、俄勒冈大学图书馆[40]、加州大学圣克鲁兹分校图书馆[41]、纽约大学图书馆[42]、拉法耶特学院图书馆[43]、麦克马斯特大学图书馆[44]、凯斯西储大学图书馆[45]、埃默里大学图书馆[46]、佐治亚州立大学图书馆[47]、哥伦比亚大学图书馆[48]、圣母大学图书馆[49]、伊利诺伊大学香槟分校图书馆[50]等15所高校图书馆为对象,主要通过访问图书馆网站、调研数字学术相关文献与网络资料的途径对其数字学术实践进行较为详细的调查(调研时间为2017年7月至9月),调查内容为各图书馆数字学术理念、空间、工具、资源、服务、人员、技术等。

调研发现,图书馆数字学术服务内容主要包括物理空间、数字科研工具、研究数据服务(另有部分图书馆未将研究数据服务纳入"数字学术服务",而是将两者作为并列的服务项目,如麦克马斯特大学图书馆)、学术交流与数字出版服务、数字人文服务、数字学术研讨与培训服务、数字基础支持服务等,图书馆员在其中扮演多样的角色、承担多样的职责。其中,研究数据服务具体包括数据管理计划咨询服务、元数据服务、数据仓储与保存服务、数据出版服务、数据分析与可视化服务等。

2.3 3所加拿大高校图书馆数据服务调研

以上述对麦克马斯特大学图书馆数字学术服务的调查为线索,同时考虑图书馆数据服务路径对笔者未来从事数据服务实践的参考作用,因此进一步选取加拿大麦克马斯特大学图书馆[51]及与其较为相似的温莎大学图书馆[52]、约克大学图书馆[53]作为深入调研对象,通过访问图书馆网站的方式了解图书馆数据服务中的数据来源、数据资源呈现方式、服务项目等内容(调研时间为2017年9月至12月)。调研发现,这3所高校图书馆的"数据服务"有两大方面:一是数据资源集成与揭示,图书馆主要对已获得授权的统计数据、研究人员产生的研究数据、开放获取的第三方数据、帮助性/学习性数据资源等进行集成,如加拿大国家统计局的部分统计数据、数据视频课程、数据与数据产品引用指南等,并按照数据学科或数据功用进行分类导航;二是数据管理与服务,如软件支持、统计咨询服务、全生命周期的研究数据管理、地图与地理空间专题数据服务等。

2.4 数据服务路径总结及其代表性图书馆

将以上调研作为基础和线索(并非严格对上述调研进行完全的解读),将高校图书馆、专业图书馆统称为"学术图书馆",秉承"空间即服务""资源即服务""平台即服务""软件即服务"等理念,本研究总结出目前图书馆较为常见的5种数据服务路径:数据平台与数据服务、"一站式"资源发现系统与数据发现、整合型数字学术空间与数据服务、开放数据资源导航与数据服务以及综合性数据咨询与支持服务。5种数据服务路径及其代表性图书馆见表1,以下将进一步对这5种数据服务路径展开阐述、例证和讨论。需说明的是,这5种路径的划分以数据服务在图书馆网站的"栏目归属"及图书馆整体服务设计为依据,不同路径可分别独立存在,也可协同发挥作用,在部分特殊情况下还可能存在交叉的部分(如在"整合型数字学术空间与数据服务"路径下提供"综合性的数据咨询与支持服务")。

表1 5种图书馆数据服务路径及其代表性图书馆

数据服务路径	代表性图书馆
数据平台与数据服务	澳大利亚联邦科学与工业研究组织图书馆、美国国家农业图书馆、德国国家科技图书馆
"一站式"资源发现系统与数据发现	加拿大联邦科学图书馆、加拿大国家科学图书馆、澳大利亚地球科学局图书馆
整合型数字学术空间与数据服务	华盛顿大学图书馆、布朗大学图书馆、俄勒冈大学图书馆、香港中文大学图书馆、纽约大学图书馆、佐治亚州立大学图书馆、哥伦比亚大学图书馆、圣母大学图书馆、伊利诺伊大学香槟分校图书馆
开放数据资源导航与数据服务	麦克马斯特大学图书馆、约克大学图书馆、温莎大学图书馆
综合性数据咨询与支持服务	德国国家医学图书馆、德国国家科技图书馆

3 5种图书馆数据服务路径及其例证

3.1 数据平台与数据服务

根据相关调研[54]可知,国内外由教学研究机构、研究资助机构、国际组织等建设的数据管理与服务平台已经较为常见,其中由图书馆组织或参与建设和维护的数据平台也逐渐丰富。目前,图书馆组织或参与的数据平台可分为两类:一类是以数据监护(data curation)生命周期为导向的开放研究数据管理平台,如约翰霍普金斯大学图书馆主导的Data Conservancy、北京大学图书馆主导的北京大学开放研究数据平台、复旦大学图书馆参与的复旦大学社会科学数据研究中心等;另一类是以机构内数据资源集成、浏览、检索为导向的数据门户,如澳大利亚联邦科学与工业研究组织图书馆主导的数据门户CSIRO Data Access Portal、德国国家科技图书馆主导的RADAR[55]。

3.1.1 研究数据管理平台　不同图书馆的数据管理平台有不同的功能侧重,如Data Conservancy强调其作

为数据监护技术基础设施的功能,而北京大学开放研究数据平台、复旦大学社会科学数据平台等强调以平台为基础的数据收集、保存、检索和利用服务。

其中,Data Conservancy 由约翰霍普金斯大学 Sheridan 图书馆始建于 2009 年,以 Fedora 系统原型为基础开发而成,受到美国国家自然科学基金会 DataNet 项目的资助,提供数据保存、共享和发现服务[56],但目前不支持数据浏览、检索和下载功能。Data Conservancy 现已开发 5 种具有元数据捕获、数据捕获/保存/可视化、数据组织与描述等功能的软件并提供免费下载使用,包括 Fedora API-X、RMap、Data Conservancy Packaging Specifications、Packaging Tool GUI、Package Ingest Service 等。

再如,北京大学开放研究数据平台于 2015 年底正式上线,由北京大学图书馆、国家自然科学基金、北京大学管理科学数据中心、北京大学科研部、北京大学社科部联合主办,北京大学图书馆负责运行维护。该平台现已集成数据空间 28 个、数据集 153 个、数据文件 533 个(截至 2018 年 1 月 31 日),覆盖社会科学、医药健康与生命科学、计算机与信息科学、商业与管理、地球与环境科学等领域[57]。

3.1.2 机构数据门户 澳大利亚联邦科学与工业研究组织(the Commonwealth Scientific and Industrial Research Organisation,CSIRO)[58]是澳大利亚最大的国家级科研机构,CSIRO 将其提供的图书馆与信息服务置于"出版物"栏目下,主要包括 CSIRO 出版、数据门户(data access portal)、研究出版物机构库及 CSIRO 图书馆服务等项目[59],为科学研究提供全方位的信息资源支撑服务。

CSIRO 数据门户[60]集成了 CISRO 机构及其研究人员发布的研究数据、软件及其他数字资产,目前已有数据和软件资源 2 000 余条(截至 2018 年 1 月 31 日),用户可浏览、检索所有数据,还可使用某些特定学科领域的检索工具(domain search tools)进行检索。

3.1.3 以数据平台为依托的数据服务可持续发展 以数据平台为依托的数据服务可持续发展问题受到国内外学者的关注,包括平台运营经费的可持续、数据资源增长压力、数据重用率提升压力以及跨机构合作机会的不确定性、研究人员需求的不确定性、数据政策的不确定性等[61]。

例如,2016 年 1 月,OCLC 针对图书馆经费紧缺、人力紧张而导致的数据管理服务困难,探讨了 7 种可能的筹资方案,包括:①获得机构预算支持;②在预算申请中加入数据监管成本;③向研究者收费;④向用户收费;⑤募集资金;⑥数据存储库资金;⑦其他可能的方案。但 OCLC 进一步针对北卡罗来纳大学 2010 位教员进行的调查显示,没有人认为应当由用户付费,而其他方案均有支持者,因此组合使用各种筹资方法可能是必须的[62]。

另有,2017 年 12 月北京大学图书馆联合北京大学信息管理系、南海大数据应用研究院等机构启动首届全国高校数据驱动创新研究大赛,赛事一经推出反响热烈[63],参赛者必须利用已经存在于北京大学开放研究数据平台的数据或自己的原创数据进行研究,并且需要将原创数据提交至该平台。该赛事有效地为平台合理征集了来自合作单位、参赛者的新的数据资源,对于促进数据资源积累、提高数据重用效率从而促进数据平台的可持续发展不失为一种有益的探索实践。

3.2 "一站式"资源发现系统与数据发现

"一站式"资源发现系统在图书馆中的应用已经比较普遍,国外如加拿大联邦科学图书馆(Federal Science Library,FSL)的"one-stop search box"、加州大学洛杉矶分校图书馆的"One Search"等,国内如清华大学图书馆的"水木搜索"、台湾大学图书馆的"整合查询系统"等。

3.2.1 国外图书馆资源发现系统与数据发现 以加拿大联邦科学图书馆[24]为例,FSL 是由加拿大农业图书馆、环境与气候变化图书馆、渔业与海洋图书馆、健康图书馆、国家科学图书馆、加拿大自然资源图书馆等 6 所不同学科领域的专业图书馆组成的图书馆资源门户,与我国的国家科技图书文献中心(NSTL)较为相似。FSL 资源发现系统以 Summon 2.0 为原型,用户可"一站式"发现 6 所图书馆所有印本馆藏和数据库资源、知识库资源和开放获取资源,也可限定只检索某一成员图书馆的资源;用户在使用 FSL 发现系统时可按照内容类型、学科分类、主题词等进行筛选。用户可检索和获取的内容类型非常多样化,包括期刊文章、图书评论、档案资料、视频记录、计算机文件、数据集、政府文件、图片、手稿、地图、微电影、新闻通知、专利、标准、演示报告、科技报告、网络资源等 40 种类型,覆盖农业、生物、化学、教育、工程等 58 个学科方向。目前,FSL 发现系统已集成了超过 3 000 万条资源记录,其中数据集 3 658 个、图片 44 张、地图 71 542 张、视频记录 2 302 个、音频记录 210 个(截至 2018 年 1 月 31 日)。

以"数据集"的发现为例,FSL 发现系统的数据集多为开放获取的网络数据资源,用户可以"Data Set"为

筛选条件浏览全部数据集或搜索某一特定主题的数据集。FSL 只收割了开放数据集的元数据信息,用户可根据 FSL 的指示链接到数据集的原始出处,再进行获取和利用。

3.2.2　国内图书馆资源发现系统与数据发现　以清华大学图书馆为例,清华大学图书馆"水木搜索"于 2011 年 12 月开通试用[64],现以 Primo 为系统原型,可供用户检索发现的资源类型包括研究数据集、统计数据、地图、乐谱、音像资料、数据库、图书、文章、期刊、图片、技术报告、法律文件、会议论文集、学位论文、专利、报纸、古籍、政府出版物、参考条目、评论等 20 余种,仅支持特定条件下的检索,不支持全部或某些内容类型资源的浏览。

以"图片"的检索发现为例,以任意字段包含"woman"为条件进行检索,获得 84 550 个结果(检索时间为 2018 年 2 月 2 日),用户可查看图片的题名、主题、描述、来源等信息,也可链接到在线资源查看图片尺寸、历史故事等更详细的描述并选择合适的图片大小进行下载。

3.2.3　数据发现功能宣传与研究　数据发现功能往往"深藏"在发现系统之中而不易被用户发现和利用,图书馆也较少对数据发现功能进行宣传,如"水木搜索"上线时清华大学图书馆介绍了其海量资源、快速检索、结果聚合、多数据源汇聚、个性化服务等特点,但未对数据(数据集、统计数据)发现功能进行推介;同时,目前针对图书馆资源发现系统的相关研究也较少提及数据发现功能,更多的是强调其海量资源、快捷获取、简单易用等方面[65]。因此,加强对图书馆资源发现系统的数据发现功能的宣传和研究是必要的。

3.3　整合型数字学术空间与数据服务

3.3.1　数字学术空间与数据服务案例　数字学术(digital scholarship)既是一种数字技术、学术产品概念,又是一种学术交流模式、教学研究范式,构成数字化、网络化、新媒体时代的数字学术环境[66],图书馆为适应环境变化和用户需求而提供数字学术服务,如华盛顿大学图书馆、加州大学圣克鲁兹分校图书馆、香港中文大学图书馆等。

不同图书馆以不同的具体名称与形式呈现其数字学术服务,如数字学术空间(digital scholarship commons)、数字学术实验室(digital scholarship LAB)、数字学术中心(centre for digital scholarship)等,既包含了物理空间、数据资源、知识服务、数字人文项目等单个元素,又是对物理空间、数据资源、知识服务、数字人文项目、研讨培训活动等系列内容的统称与整合,研究数据服务是数字学术服务的重要组成部分。

"数字学术空间"框架内的研究数据服务基本覆盖了数据生命周期的各个环节,尤以咨询服务和数据分析与可视化服务最为突出:①咨询服务包括数据管理计划提纲、指南、最佳实践咨询以及元数据方案、标准咨询等,如俄勒冈大学图书馆为用户提供数据管理计划提纲[67](如何进行数据描述、元数据及工具选择、数据组织/存储/保护、数据共享、数据长期保存等内容)和数据管理计划最佳实践/指南[68](如视频数据、地理空间数据、图片数据、文本数据最常用的标准格式推荐)。②数据分析与可视化服务包括数据分析软件(如 ATLAS.ti、GRASS GIS、SPSS)的提供或推荐和可视化设备(如高分辨率显示屏、交互墙)及对应场地的提供,如香港中文大学图书馆提供地理信息系统软件 GRASS GIS、统计分析软件 IBM SPSS、统计计算及图像软件 R、影像化及分析医学数据软件 TreeAge Pro Healthcare 等[69],布朗大学洛克菲勒图书馆的数字学术实验室配备了 7x16 英尺的高分辨率显示墙[70]。

3.3.2　数据服务的关联协同因素　将前文所述的数字学术空间下的数据服务与数据管理平台导向的数据服务进行对比可知,通常情况下前者强调数据资源本身驱动的数据服务,而后者除数据资源外还强调辅助资源(软件、硬件、空间、技术、标准、指南等)驱动的数据服务。由此可以认为,软件、硬件、空间、技术等是数据服务的关联协同因素,能够实现数据可视化、数据交互、数字人文、数据技能培训、数据驱动的研究成果展示等功能,使得数据服务内容更加丰富、数据服务体系更加完善的同时更好地提升用户体验、满足用户需求。

3.4　开放数据资源导航与数据服务

3.4.1　国内外图书馆数据资源导航服务案例　开放获取运动深入发展,开放创新、开放科学、开放数据正在成为学术界的共识。随着数据开放共享政策、标准、技术等的不断发展完善以及开放数据资源数量的不断积累,开放数据成为数据资源不可或缺的组成部分。目前,除部分图书馆"一站式"资源发现系统收割开放数据资源的元数据之外,针对开放数据知识库进行数据资源导航也成为数据服务的一种形式,如加拿大约克大学图书馆、温莎大学图书馆和中国科学院文献情报中心等。

国外开放数据资源导航服务以约克大学图书馆为例,该图书馆主要对来自加拿大、美国及国际组织等的开放数据进行导航[71],为方便读者准确查找所需数

据,约克大学图书馆按照学科、功用、地域等对数据资源进行分类导航(提供数据资源原始链接,而非直接从约克大学图书馆下载),如艺术、文化与娱乐,人口普查,选举与政治,人权,女性,公共部门等。

国内开放数据资源导航服务以中国科学院文献情报中心为例,该中心除对其馆藏数据资源进行推介之外,还对大量的数据平台、数据期刊、官方统计局、国际组织等的开放数据进行集成导航[72],主要包括:①国家科技基础条件平台的20余个数据平台,如国家地震科学数据共享中心、国家农业科学数据共享中心、国家人口与健康科学数据共享服务平台等;②中国科学院近30个科学数据库,如空间科学数据中心、材料科学领域基础科学数据库、聚变数据库等;③国内外知名通用型数据知识库,如figshare、Dryad、GenBank等;④国内外开放获取数据期刊,如*Scientific Data*、*Earth System Science Data*、《中国科学数据》等;⑤国际组织开放数据平台,如世界银行开放数据(World Bank Open Data)、国际货币基金组织数据(IMF Data)、经济合作与发展组织数据平台(OECD)等;⑥世界部分主要国家统计局,如加拿大国家统计局、新加坡国家统计局、中国国家统计局等。

3.4.2 数据资源导航的数据粒度与质量　对比发现,约克大学图书馆和中国科学院文献情报中心的数据分类导航是两种不同的思路,前者按照学科、功用等的混合方向,后者按照常见的开放数据获取途径和粗略的学科门类,前者分类粒度更小而后者数据揭示相对粗放。若后者以更小粒度(如数据期刊内的单篇数据论文、数据库中的单个数据集或单个数据文件)进行导航,在资源不断增长、导航不断更新的情况下又容易显得"臃肿"和"冗余";同时,开放数据资源的海量增长提高了数据质量控制的难度,因而相应地提高了数据导航对高质量数据资源选择的难度,因此数据资源导航的数据粒度、质量控制问题还有待进一步研究。

3.5 综合性数据咨询与支持服务

根据多份对图书馆研究数据服务的调查报告[3-4,73]可知,目前图书馆提供的技术/实践类数据服务较少,而以咨询/支持类数据服务居多,事实证明确实如此。

3.5.1 数据管理咨询与支持

(1)"自问自答型"咨询。为便于用户更好地理解和接受数据管理的相关要求,图书馆通过"自问自答"的方式,向用户提供关于数据管理必要性与重要性、数据管理计划及其工具、数据政策、数据文件格式、数据存储、数据引用等的咨询与引导。例如,哈佛大学定量社会科学研究所(Institute for Quantitative Social Science,IQSS)组织建设了数据平台Harvard Dataverse并进行管理维护,而哈佛大学图书馆则围绕数据管理计划(DMP)及其工具、数据政策与指南、数据文件格式与命名、数据存储、数据引用等提供咨询与帮助性资源支持。其中,针对"数据管理计划"提出并回答"为什么要管理数据""什么是数据管理计划""我的数据管理计划应包括什么""数据管理计划样例""哈佛大学图书馆在数据管理中扮演的角色"等问题[74];针对"数据管理计划工具"提供"什么是DMPTool""在哈佛使用DMPTool的途径""使用DMPTool的好处(视频)"的帮助性资源支持[75]。

(2)覆盖生命周期的咨询。即便只有部分图书馆参与了数据管理平台的建设或维护,但图书馆围绕数据管理生命周期提供教育类、专业咨询类数据服务依然成为了一种越来越普遍的现象。

例如,麦克马斯特大学图书馆覆盖数据管理生命周期的咨询服务主要包括[76]:①数据计划 & 准备阶段,包括:数据管理自我测试(提问形式);数据管理计划释义、工具及模板推荐;加拿大主要研究资助机构对研究数据管理的要求等。②数据收集 & 分析阶段,包括:解释元数据的作用、提供通用型和学科专用型元数据标准的简介、推荐元数据标准及最佳实践;数据存储与安全的建议与注意事项;数据格式化及数据格式转换建议。③数据保存 & 归档三大阶段,包括:数据归档及长期保存位置与最佳实践指南;数据共享的益处及相关规定与限制;数据知识产权与数据敏感性因素等。

再如,约克大学图书馆从"开始研究前""研究过程中""研究完成后"3个阶段,引导和帮助研究人员管理研究数据[77]:①在"开始研究前",研究人员应了解研究资助机构的数据政策;遵循研究伦理;创建研究数据管理计划(引导式问题)。②在"研究过程中",研究人员应注意数据存储、共享、安全和保存问题(引导式提问);应将研究数据文档化(引导式提问);如何发现和引用研究数据(推荐数据知识库列表和数据知识库注册系统以便发现数据、提供数据引用指南及样例以便规范引用数据)。③在"研究完成后",研究人员需进行数据存储(提供本校学者库、机构知识库等存储途径);明确数据许可(列举常见的许可协议,如CC0、CC-BY)。

3.5.2 其他数据咨询与支持服务　除数据管理咨询

外,图书馆还提供其他多样的数据咨询与支持服务,包括 DOI 注册、数据出版、研讨与培训等。例如,在 DOI 分配方面,德国国家科技图书馆负责为技术与工程、建筑、化学、计算机科学、数学、物理学等领域的研究数据及其他研究产品提供 DOI 注册服务[78];加拿大国家科学图书馆及其所属母体机构加拿大国家研究理事会是 DataCite 的创始成员并创建了加拿大的 Datacite(Data-Cite Canada),负责与加拿大的各类数据中心、图书馆开展广泛合作,以提供研究数据 DOI 服务[79],致力于提高数据的可发现、可访问和可重用。再如,在数据出版方面,当相应研究人员在《德国医学科学》(*German Medical Science*,GMS)期刊发表文章时,德国国家医学图书馆(ZB MED)负责在数据知识库 Dryad 中存储并出版文章对应的研究数据并建立两者之间的链接,并承担研究数据出版的成本费用;另外,ZB MED 提供两种独立数据出版的方式:一是研究人员可将各种格式的数据单独或连同文章一起在专家库中存储和出版;二是特定学科数据知识库,在农业、生物、环境、医药学、神经科学、营养学等特定学科数据知识库中存储和出版对应的研究数据[80]。

3.5.3 数据咨询与支持服务的差异化策略　前文调研发现,国外围绕研究数据管理的咨询较多而国内则较少,政策导向是其中的重要原因。例如,2003 年美国国立卫生研究院开始要求(希望)受到超过 50 万美元的资助者提交数据共享计划[81],2011 年 1 月美国国家科学基金会开始要求申请资助者提交数据管理计划[82],此后美国研究型大学逐渐开始强调数据管理,图书馆也根据自身情况(如伊利诺伊大学香槟分校认为其产生了许多出版物和数据集,但却没有做多少数据管理工作,无形中助长了该校从事数据密集型研究的障碍)逐渐承担数据管理职责[61]。然而,目前我国以国家自然科学基金、国家社会科学基金、中国科学院等为代表的研究资助机构尚未有明确的数据管理要求,因此图书馆面向用户的数据管理咨询服务可能暂不适用,反而是开放数据获取、数据权益、数据许可、数据知识产权、数据驱动研究的经典案例等内容可能更受欢迎。因此,针对不同学术机制、不同数据政策、不同学科特点下的数据咨询与支持服务进行差异化探索,将有助于更加精准地提供数据服务、满足用户需求。

4 结语

前文所述的学术图书馆数据服务的 5 种路径各有特点和适用性。

显然,5 种路径既可独立存在、又可协同发挥作用,如图书馆可在运营数据管理平台的同时提供综合性数据咨询与支持服务(如北京大学图书馆),既可进行开放数据资源导航又可提供综合性数据咨询与支持服务(如约克大学图书馆)。

事实上,不同的服务路径各有其优劣或适用性条件,如综合性数据咨询与支持服务的“门槛”较低,最为方便快捷,但纯粹的咨询与支持服务容易产生“粗浅”和“表面”的缺点,能够持续产生的效果可能比较有限;而数据平台建设的“门槛”较高,需要充足的资金、技术、经费、政策力量的支持,还可能面临数据增长、更新缓慢甚至难以可持续的压力并且可能因数据重用率过低而陷入数据“休眠”的困窘状态;部分图书馆“一站式”资源发现系统并未覆盖网络资源的发现,发现系统对网络资源(开放数据资源)的覆盖需要合适的系统建设或改造的契机,也倚赖强大的技术力量;整合型数字学术服务体系内的数据服务需与空间、软件、硬件等其他因素协同才能发挥更大的作用;开放数据资源导航服务的可操作性较强、“门槛”也较低,但图书馆仅在其中发挥“中介”的作用,较难形成核心竞争力和用户粘性。

然而,5 种数据服务模式与路径都有其存在的价值和必要性,任意一种数据服务模式的存在都有助于构建完整的数据服务体系,帮助图书馆在与出版商、信息技术与服务机构等的竞争与合作中更好地占据数据服务的“领地”,把握数据服务的主动权。

参考文献:

[1] 司莉，庄晓喆，王思敏，等. 2005 年以来国外科学数据管理与共享研究进展与启示[J]. 国家图书馆学刊，2013，22(3):40－49.

[2] NELSON M S. Connecting with data: first steps toward an emerging area of library service[J]. American Society for Engineering Education，2011,375(22):1－17.

[3] TENOPIR C，BIRCH B，ALLARD S. Academic libraries and research data services: current practices and plans for the future[R/OL]. [2018－01－09]. http://www.ala.org/acrl/sites/ala.org.acrl/files/content/publications/whitepapers/Tenopir_Birch_Allard.pdf.

[4] TENOPIR C，TALJA S，HORSTMANN W，et al. Research data services in European academic research libraries[EB/OL]. [2018－01－09]. http://libereurope.eu/wp-content/uploads/2016/10/LIBER_RDM_SURVEY_FINAL_MANUSCRIPT.pdf.

[5] 安艳杰，黄良军. 图书馆参考咨询新领域——数据服务[J]. 情报资料工作，2008(3):94－96.

[6] 陈建新. 科学数据服务:图书馆服务的新领域[J]. 图书与情报, 2013(4):93-95.
[7] 沈婷婷. 北美科学数据服务的实践与启示[J]. 现代情报, 2014, 34(12):145-147.
[8] 张颖. 美国研究型图书馆研究数据服务的实践进展及趋势[J]. 图书情报工作, 2017,61(9):33-41.
[9] 陈媛媛, 柯平.《欧洲学术研究图书馆科研数据服务》调查对我国高校图书馆科研数据服务的启示[J]. 图书情报工作, 2017, 61(6):73-78.
[10] 柳影, 吴鸣, 胡卉, 等. 国外高校图书馆科研数据服务研究——以美国伊利诺伊大学香槟分校为例[J]. 新世纪图书馆, 2017(6):69-74.
[11] 黄鑫, 邓仲华. 国外高校图书馆科学数据的元数据服务研究[J]. 图书与情报, 2017(2):84-90.
[12] 解金兰, 冉从敬. 图书馆数据服务生态圈的架构思路与功能研究[J]. 图书馆理论与实践, 2017(6):94-98.
[13] 黄鑫, 邓仲华. "互联网+"视角下的图书馆科学数据服务模型[J]. 情报理论与实践, 2017, 40(3):75-80.
[14] 朱彩萍. 高校图书馆提供科学数据服务的途径与内容[J]. 图书与情报, 2014(3):97-99.
[15] 张凯勇. 数据密集型科学环境下的高校图书馆科学数据服务[J]. 图书馆学研究, 2014(3):69-72.
[16] 赵欣. 高校图书馆开放数据服务平台研究[J]. 兰台世界, 2016(3):69-70.
[17] 徐刘靖, 沈婷婷. 高校图书馆员数据素养内涵及培养机制研究[J]. 图书馆建设, 2016(5):89-94.
[18] 蒋丽丽, 陈幼华, 陈琛. 国外高校图书馆数据馆员服务模式研究[J]. 图书情报工作, 2015, 59(17):56-61.
[19] German National Library of Science and Technology[EB/OL]. [2017-07-25]. https://www.tib.eu/en/.
[20] The German National Library of Medicine[EB/OL]. [2017-07-26]. https://www.zbmed.de/en/.
[21] National Library of Medicine[EB/OL]. [2017-07-26]. https://www.nlm.nih.gov/.
[22] National Agricultural Library[EB/OL]. [2017-12-18]. https://www.nal.usda.gov/.
[23] NIST Research Library[EB/OL]. [2017-08-06]. https://www.nist.gov/nist-research-library.
[24] Federal Science Library[EB/OL]. [2017-08-07]. http://science-libraries.canada.ca/eng/home/.
[25] National Science Library[EB/OL]. [2017-08-07]. http://science-libraries.canada.ca/eng/national-science-library/.
[26] CSIRO Libraries [EB/OL]. [2017-12-14]. https://www.csiro.au/en/Publications/Libraries.
[27] Fisher Geoscience Library[EB/OL]. [2017-12-16]. http://www.ga.gov.au/data-pubs/library.
[28] National Institute of Informatics Library[EB/OL]. [2018-02-06]. http://www.nii.ac.jp/en/about/library/.
[29] JAEA Library[EB/OL]. [2017-08-08]. http://tenkai.jaea.go.jp/english/library/.
[30] Russian National Public Library for Science and Technology[EB/OL]. [2017-08-06]. http://english.gpntb.ru/.
[31] National Physical Laboratory Library[EB/OL]. [2017-08-09]. http://www.npl.co.uk/library/.
[32] 中国台湾"中央研究院"图书馆服务[EB/OL]. [2017-08-04]. http://aslib.sinica.edu.tw/.
[33] 介凤, 盛兴军. 数字学术中心:图书馆服务转型与空间变革——以北美地区大学图书馆为例[J]. 图书情报工作, 2016, 60(13):64-70.
[34] 李红培. 国外高校图书馆数字学术空间的实践及启示[J]. 图书馆理论与实践, 2016(7):89-92.
[35] BRYAN S. The university library as incubator for digital scholarship[EB/OL]. [2017-07-20]. http://er.educause.edu/articles/2014/6/the-university-library-as-incubator-for-digital-scholarship.
[36] 香港中文大学图书馆. 数码学术研究服务[EB/OL]. [2017-07-18]. http://www.lib.cuhk.edu.hk/sc/research/digital-scholarship.
[37] Libraries of University of Washington. Digital Scholarship[EB/OL]. [2017-07-19]. http://www.lib.washington.edu/digitalscholarship/.
[38] Brown Library. Center for Digital Scholarship[EB/OL]. [2017-07-19]. https://library.brown.edu/cds/.
[39] University of Richmond. Digital Scholarship LAB[EB/OL]. [2017-07-20]. http://dsl.richmond.edu/.
[40] University of Oregon Libraries. Digital Scholarship Center[EB/OL]. [2017-07-20]. https://library.uoregon.edu/digital-scholarship-center.
[41] UC Santa Cruz University Library. David Kirk digital scholarship commons[EB/OL]. [2017-07-30]. https://guides.library.ucsc.edu/DSC.
[42] New York University Library[EB/OL]. [2017-08-01]. https://library.nyu.edu/departments/digital-scholarship-services/.
[43] Lafayette College Library. Digital scholarship services[EB/OL]. [2017-07-20]. http://digital.lafayette.edu/.
[44] McMaster University Library. Sherman Center for Digital Scholarship[EB/OL]. [2017-07-20]. http://scds.ca/what-we-do/.
[45] Case Western Reserve University Kelvin Smith Library. The Freedman Center for Digital Scholarship[EB/OL]. [2017-07-21]. http://library.case.edu/ksl/freedmancenter/.
[46] Emory Libraries &Information Technology. Emory Center for Digital Scholarship[EB/OL]. [2017-08-15]. http://digitalscholarship.emory.edu/.
[47] Georgia State University Library. CURVE[EB/OL]. [2017-07-21]. http://research.library.gsu.edu/curve.
[48] Columbia University Libraries Center for Digital Research and Scholarship[EB/OL]. [2017-07-21]. http://cdrs.columbia.edu/cdrsmain/.

[49] University of Notre Dame. Center for Digital Scholarship[EB/OL].[2017-07-21]. http://library.nd.edu/cds/.
[50] University library of Illinois at Urbana-Champaign. Scholarly Commons[EB/OL].[2017-07-21]. http://www.library.illinois.edu/sc/.
[51] McMaster University Library[EB/OL].[2017-10-06]. http://library.mcmaster.ca/.
[52] University of Windsor Leddy Library. Academic Library Centre [EB/OL].[2017-09-26]. http://leddy.uwindsor.ca/adc.
[53] York University Library[EB/OL].[2017-09-30]. https://www.library.yorku.ca/web/.
[54] 卫军朝,张春芳.国内外科学数据管理平台比较研究[J].图书情报知识,2017(5):97-107.
[55] RADAR[EB/OL].[2017-08-30]. https://www.radar-service.eu/en.
[56] Data Conservancy[EB/OL].[2018-01-31]. http://dataconservancy.org/about/history-of-dc/.
[57] 北京大学开放研究数据平台[EB/OL].[2018-01-31]. http://opendata.pku.edu.cn/dataverse/pku.
[58] CSIRO[EB/OL].[2017-08-04]. https://www.csiro.au/.
[59] CSIRO. Publications[EB/OL].[2017-12-14]. https://www.csiro.au/en/Publications.
[60] CSIRO Data Access Portal[EB/OL].[2017-12-14]. https://data.csiro.au/dap/home? execution=e4s1.
[61] BRYANT R, LAVOIE B, MALPAS C. The realities of research data management——part two: scoping the university RDM service bundle[EB/OL].[2018-02-10]. https://www.oclc.org/content/dam/research/publications/2017/oclcresearch-rdm-part-two-scoping-service-model-2017.pdf.
[62] ERWAY R, RINEHART A. If you build it, will they fund? Making research data management sustainable[EB/OL].[2018-02-09]. http://www.oclc.org/content/dam/research/publications/2016/oclcresearch-making-research-datamanagement-sustainable-2016.pdf.
[63] 北京大学图书馆."首届全国高校数据驱动创新研究大赛"奖项名额增加[EB/OL].[2018-01-31]. http://www.lib.pku.edu.cn/portal/cn/news/0000001644.
[64] 清华大学图书馆公告消息. 清华大学学术信息发现平台——"水木搜索"开通试用[EB/OL].[2018-02-10]. http://lib.tsinghua.edu.cn/dra/news/annoucement/3007.
[65] 李雯,龙旭梅.图书馆主流资源发现系统的现状和未来展望[J].中华医学图书情报杂志,2016,25(5):26-29.
[66] 刘兹恒,涂志芳.数字学术环境下学术图书馆发展新形态研究——以空间、资源和服务"三要素"为视角[J].图书情报工作,2017,61(16):15-23.
[67] University of Oregon Libraries. Creating a data management plan [EB/OL].[2017-07-20]. https://library.uoregon.edu/datamanagement/dmp.
[68] University of Oregon Libraries. Data management plan best practices/guidelines[EB/OL].[2017-07-20]. https://library.uoregon.edu/datamanagement/guidelines.
[69] 香港中文大学图书馆.数码学术研究服务-工具[EB/OL].[2017-07-20]. http://www.lib.cuhk.edu.hk/sc/research/digital-scholarship/tools.
[70] Brown Library. Center for Digital Scholarship-activities[EB/OL].[2017-07-20]. http://library.brown.edu/cds/activities/.
[71] York University Library. Data & Statistics[EB/OL].[2017-09-30]. http://researchguides.library.yorku.ca/data? hs=a&gid=1377.
[72] 中国科学院文献情报中心.数据服务[EB/OL].[2018-02-02]. http://dds.las.ac.cn/rds.html.
[73] BRYANT R, LAVOIE B, MALPAS C. The realities of research data management——part one: a tour of the research data management (RDM) service space[R/OL].[2018-01-26]. http://www.oclc.org/content/dam/research/publications/2017/oclcresearch-research-data-management-service-space-tour-2017.pdf.
[74] Harvard Library. Data management plan[EB/OL].[2018-02-02]. https://guides.library.harvard.edu/dmp.
[75] Harvard Library. Data management - DMPTool[EB/OL].[2018-02-02]. https://guides.library.harvard.edu/c.php? g=471243&p=3223151.
[76] McMaster University Library. Research data management at McMaster[EB/OL].[2017-09-28]. http://library.mcmaster.ca/rdm.
[77] York University Library. Research data management overview[EB/OL].[2017-09-30]. https://www.library.yorku.ca/web/open/overview/
[78] German National Library of Science and Technology. DOI registration[EB/OL].[2017-08-30]. https://www.tib.eu/en/publishing-archiving/doi-service/doi-registration/.
[79] DataCite. Members[EB/OL].[2017-12-20]. https://www.datacite.org/members.html.
[80] ZB MED. Publishing research data[EB/OL].[2017-07-27]. https://www.publisso.de/en/publishing/publishing-research-data/.
[81] National Institutes of Health (NIH). Final NIH statement on sharing research data[EB/OL].[2018-02-10]. https://grants.nih.gov/grants/guide/notice-files/NOT-OD-03-032.html.
[82] National Science Foundation. Dissemination and sharing of research results[EB/OL].[2018-02-10]. https://www.nsf.gov/bfa/dias/policy/dmp.jsp.

作者贡献说明:

涂志芳:负责网络调研、文献调研、论文撰写与修改;
徐慧芳:负责论文修改与审阅。

Study on the Multi-path of Research Data Services in Academic Libraries

Tu Zhifang Xu Huifang
National Science Library, Chinese Academy of Sciences, Beijing 100190

Abstract: [**Purpose/significance**] Studying on the existing research data services paths of academic libraries can provide reference for domestic libraries to explore the path and model of data services. [**Method/process**] Based on the previous researches on 14 professional libraries' services, 15 university libraries' digital scholarship services, and 3 Canada university libraries' data services, the paper summarizes five main paths of data services: data platforms and data services, "one-stop" resource discovery systems and data discovery, digital scholarship services and data services, open data navigation and data services, and comprehensive data consulting and supporting services. Centering on the five path, the paper analyzes their features, makes case studies and discussion on related issues and solutions that require further study. [**Result/conclusion**] Different data services paths have advantages and disadvantages and applicability conditions, and they can work independently and collaboratively as well. The five paths have their own value and necessity, any one of them contributes to construct a complete data service system, and therefore enables library to better occupy the "domain" of data services and to take the initiative in data services in competition and cooperation with publishers, information technology and service organizations.

Keywords: academic library research data services data management platform resource discovery system digital scholarship